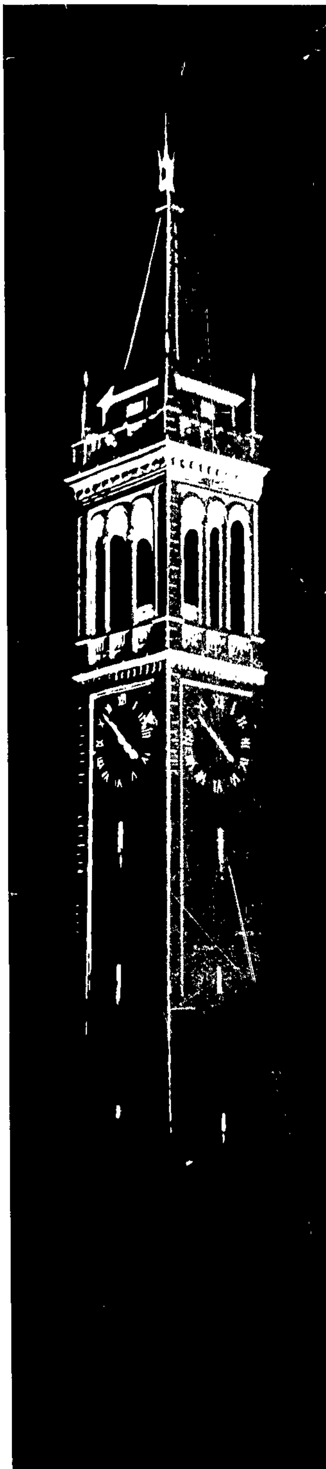

Discrete Dynamic Programming

by

D. Blackwell

Series No. 60, Issue No. 468
Contract No. Nonr-222(53)
August 13, 1962

ELECTRONICS RESEARCH LABORATORY
UNIVERSITY OF CALIFORNIA
BERKELEY, CALIFORNIA

Electronics Research Laboratory
University of California
Berkeley, California

DISCRETE DYNAMIC PROGRAMMING

by

D. Blackwell

Institute of Engineering Research
Series No. 60, Issue No. 468

Contract No. Nonr-222(53)
August 13, 1962

Reprinted from THE ANNALS OF MATHEMATICAL STATISTICS
Vol. 33, No. 2, June, 1962
Printed in U.S.A.

DISCRETE DYNAMIC PROGRAMMING[1]

BY DAVID BLACKWELL

University of California, Berkeley

1. Introduction and summary. We consider a system with a finite number S of states s, labeled by the integers $1, 2, \cdots, S$. Periodically, say once a day, we observe the current state of the system, and then choose an action a from a finite set A of possible actions. As a joint result of the current state s and the chosen action a, two things happen: (1) we receive an immediate income $i(s, a)$ and (2) the system moves to a new state s' with the probability of a particular new state s' given by a function $q = q(s' \mid s, a)$. Finally there is specified a discount factor β, $0 \leqq \beta < 1$, so that the value of unit income n days in the future is β^n. Our problem is to choose a policy which maximizes our total expected income. This problem, which is an interesting special case of the general dynamic programming problem, has been solved by Howard in his excellent book [3]. The case $\beta = 1$, also studied by Howard, is substantially more difficult. We shall obtain in this case results slightly beyond those of Howard, though still not complete. Our method, which treats $\beta = 1$ as a limiting case of $\beta < 1$, seems rather simpler than Howard's.

2. Definitions and notation. Denote by F the (finite) set of functions f from S to A. By a *policy* π, we mean a sequence $\{f_n, n = 1, 2, \cdots\}$ of functions $f_n \varepsilon F$. Using policy π means that, if we find the system in state s on the nth day, the action chosen that day is $f_n(s)$. For any sequence $g_1, \cdots, g_N$, $g_n \varepsilon F$, and any policy $\pi = \{f_n\}$, we denote by $g_1, \cdots, g_N, \pi$ the policy $\{h_n\}$ with $h_n = g_n$, $1 \leqq n \leqq N$, $h_n = f_{n-N}$, $n > N$. For any $g \varepsilon F$, we denote by $g^{(N)}, \pi$ the policy $\{h_n\}$ with $h_n = g$, $1 \leqq n \leqq N$, $h_n = f_{n-N}$, $n > N$, and by $g^{(\infty)}$ the policy $\{h_n\}$ with $h_n = g$ for all n. Finally, we denote by $T\pi$ the policy $\{h_n\}$ with $h_n = f_{n+1}$ for all n.

We associate with each $f \varepsilon F$ (1) the $S \times 1$ column vector $r(f)$ whose sth element is $i(s, f(s))$, and (2) the $S \times S$ Markov matrix $Q(f)$ whose (s, s') element is $q(s' \mid s, f(s))$. Thus $r(f)$ and $Q(f)$ specify the income and the law of motion, as a function of the current state, on a day when our rule of action is f. If we use policy $\pi = \{f_n\}$ and the system is initially in state s, the probability that the system will be in state s' at the end of the nth day is the (s, s') element of the matrix $Q_n(\pi) = Q(f_1)Q(f_2)\cdots Q(f_n)$. Thus the total expected return from π is the column vector

$$V(\pi) = \sum_{n=0}^{\infty} \beta^n Q_n(\pi) r(f_{n+1}),$$

Received September 22, 1961.

[1] This research was supported by the Information Systems Branch of the Office of Naval Research under Contract Nonr 222(53).

where $Q_0(\pi) = I$, the $S \times S$ identity matrix. We have

$$V(\pi) = r(f_1) + \beta Q(f_1) \sum_{n=1}^{\infty} Q_{n-1}(T\pi) r(f_{n+1})$$

$$= r(f_1) + \beta Q(f_1) V(T\pi).$$

We associate with each $f \varepsilon F$ the transformation $L(f)$ which maps the $S \times 1$ column vector w into $L(f)w = r(f) + \beta Q(f)w$. Thus $V(f, \pi) = L(f)V(\pi)$, and $V(f_1, \cdots, f_N, \pi) = L(f_1) \cdots L(f_N)V(\pi)$. For any two column vectors w_1, w_2, we write $w_1 \geqq w_2$ if every coordinate of w_1 is at least as large as the corresponding coordinate of w_2, and $w_1 > w_2$ if $w_1 \geqq w_2$ and $w_1 \neq w_2$. Note that $L(f)$ is *monotone*, i.e., $w_1 \geqq w_2$ implies $L(f)w_1 \geqq L(f)w_2$.

For any two policies π_1, π_2, we write $\pi_1 \geqq \pi_2$ if $V(\pi_1) \geqq V(\pi_2)$, and $\pi_1 > \pi_2$ if $V(\pi_1) > V(\pi_2)$. A policy π^* is called *optimal* if $\pi^* \geqq \pi$ for all π.

3. Optimal policies for $\beta < 1$. The methods of this section are familiar to workers in dynamic programming, from the work of Dvoretzky, Kiefer, and Wolfowitz [2], Karlin [4], and Bellman [1].

THEOREM 1. *If $\pi^* \geqq (f, \pi^*)$ for all $f \varepsilon F$, then π^* is optimal.*

PROOF. Our hypothesis is that

$$L(f)V(\pi^*) \leqq V(\pi^*) \qquad \text{for all } f \varepsilon F.$$

Then for any policy $\pi = \{f_n\}$, we have $L(f_N)V(\pi^*) \leqq V(\pi^*)$, so that, using the monotoneity of $L(f_1) \cdots L(f_{N-1})$, $L(f_1) \cdots L(f_N)V(\pi^*) \leqq L(f_1) \cdots L(f_{N-1})V(\pi^*)$, i.e., $(f_1, \cdots, f_N, \pi^*) \leqq (f_1, \cdots, f_{N-1}, \pi^*)$. Thus

$$\pi^* \geqq (f_1, \cdots, f_N, \pi^*)$$

for all N, i.e., $V(\pi^*) \geqq V(f_1, \cdots, f_N, \pi^*)$ for all N. Letting $N \to \infty$ we obtain $(\beta < 1)$,

$$V(\pi^*) \geqq V(\pi),$$

and the proof is complete.

THEOREM 2. *If $(f, \pi) > \pi$, then $f^{(\infty)} > \pi$.*

PROOF. Our hypothesis is $L(f)V(\pi) > V(\pi)$. Applying the monotone operator $L^{N-1}(f)$ yields

$$L^N(f)V(\pi) \geqq L^{N-1}(f)V(\pi),$$

so that $(f^{(N)}, \pi) \geqq (f, \pi)$ for all $N \geqq 1$. Letting $N \to \infty$ yields $f^{(\infty)} \geqq (f, \pi)$, so that $f^{(\infty)} > \pi$.

Our principal result, describing the Howard policy improvement routine for $\beta < 1$, is

THEOREM 3. *Take any $f \varepsilon F$. For each $s \varepsilon S$ denote by $G(s, f)$ the set of all a for which*

$$i(s, a) + \beta p(s, a)V(f^{(\infty)}) > V_s(f^{(\infty)}),$$

where $p(s, a)$ *is the* $1 \times S$ *row vector whose* s'th *coordinate is* $q(s' \mid s, a)$ *and* $V_s(f^{(\infty)})$ *denotes the* sth *coordinate of* $V(f^{(\infty)})$. *If* $G(s, f)$ *is empty for all* s, *then* $f^{(\infty)}$ *is optimal. For any* g *such that*

(a) $g(s) \in G(s, f)$ *for some* s *and*

(b) $g(s) = f(s)$ *whenever* $g(s) \not\in G(s, f)$, *we have* $g^{(\infty)} > f^{(\infty)}$.

PROOF. The sth coordinate of $V(g, f^{(\infty)})$ is $i(s, g(s)) + \beta p(s, g(s)) V(f^{(\infty)})$. This will exceed $V_s(f^{(\infty)})$ if and only if $g(s) \in G(s, f)$, and will equal $V_s(f^{(\infty)})$ if $g(s) = f(s)$. Thus if $G(s, f)$ is empty for all s, $f^{(\infty)} \geqq (g, f^{(\infty)})$, for all g so that, from Theorem 1, $f^{(\infty)}$ is optimal. On the other hand, for any g satisfying (a) and (b), we have $(g, f^{(\infty)}) > f^{(\infty)}$ so that, from Theorem 2, $g^{(\infty)} > f^{(\infty)}$.

Call a policy $\pi = \{f_n\}$ *stationary* if f_n is independent of n, i.e., if $\pi = f^{(\infty)}$ for some $f \in F$. As a consequence of Theorem 3, we have the

COROLLARY. *There is an optimal policy which is stationary.*

PROOF. According to Theorem 3, if we take any stationary policy $f^{(\infty)}$, either it is optimal (case $G(s, f)$ empty for all s) or it has a stationary improvement $g^{(\infty)}$ (case $G(s, f)$ nonempty for some s). Since there are only finitely many stationary policies, there is one which has no stationary improvement, so that it must be optimal.

4. Optimal policies for $\beta = 1$. For the case $\beta = 1$, the total income from a given policy is typically infinite. We may attempt instead to maximize the average rate of income or to find policies which are optimal for all β sufficiently near 1. We shall adopt the second approach. Since β is now variable, it will sometimes be desirable to exhibit the dependence of $V(\pi)$ and other quantities on β; thus we shall write $V_\beta(\pi)$ and speak of β-optimal policies. Denote by $U(\beta)$ the expected total return from a β-optimal policy. We shall say that a policy π is *optimal* if it is β-optimal for all β sufficiently near 1, i.e., if $V_\beta(\pi) = U(\beta)$ for all β sufficiently near 1, and shall say that π is *nearly optimal* if

$$U(\beta) - V_\beta(\pi) \to 0 \qquad \text{as } \beta \to 1.$$

Our problem is then to find optimal and nearly optimal policies.

We shall need certain known facts about Markov matrices, summarized as

LEMMA 1. *Let* Q *be any* $S \times S$ *Markov matrix.*

(a) *The sequence* $I + Q + \cdots + Q^N/N + 1$ *converges as* $N \to \infty$ *to a Markov matrix* Q^* *such that*

$$QQ^* = Q^*Q = Q^*Q^* = Q^*,$$

(b) *rank* $(I - Q) +$ *rank* $Q^* = S$.

(c) *For every* $S \times 1$ *column vector* c, *the system*

$$Qx = x, \qquad Q^*x = Q^*c$$

has a unique solution.

(d) $I - (Q - Q^*)$ *is nonsingular, and*

$$H(\beta) = \sum_0^\infty \beta^n(Q^n - Q^*) \to H = (I - Q + Q^*)^{-1} - Q^*$$

as $\beta \to 1$.

$$H(\beta)Q^* = Q^*H(\beta) = HQ^* = Q^*H = 0$$

and

$$(I - Q)H = H(I - Q) = I - Q^*.$$

These facts may all be found in Kemeny and Snell [5]; we indicate the proof of (d) only.

PROOF OF (d). From (a) we have, for $n > 0$, $Q^n - Q^* = (Q - Q^*)^n$, so that $H(\beta) = \sum_0^\infty \beta^n(Q - Q^*)^n - Q^* = [I - \beta(Q - Q^*)]^{-1} - Q^*$, i.e.,

$$(H(\beta) + Q^*)(I - \beta(Q - Q^*)) = I,$$

i.e.,

$$(H(\beta) + Q^*)(I - Q + Q^*) = I - (1 - \beta)H(\beta)(Q - Q^*). \tag{1}$$

Now $C - 1$ summability of $\{Q^n\}$ to Q^* implies Abel summability of $\{Q^n - Q^*\}$ to Q:

$$(1 - \beta)\sum_0^\infty \beta^n(Q^n - Q^*) = (1 - \beta)H(\beta) \to 0 \qquad \text{as } \beta \to 1.$$

Thus the matrix on the right of (1) goes to I as $\beta \to 1$, and $I - Q + Q^*$ is nonsingular. Multiplying (1) by $(I - Q + Q^*)^{-1}$ and letting $\beta \to 1$ yields $H(\beta) + Q^* \to (I - Q + Q^*)^{-1}$ as $\beta \to 1$. Verification of the equalities asserted in (d) is straightforward.

Our results for $\beta = 1$ are summarized as Theorem 4 below. We shall sometimes, to simplify statements, speak of "the policy f" when we mean the policy $f^{(\infty)}$. For example, we write $V_\beta(f)$ instead of $V_\beta(f^{(\infty)})$.

THEOREM 4. *Take any* $f \in F$ *and denote by* $Q^*(f)$ *the matrix* Q^* *associated with* $Q(f)$. *Then*

(a) $$V_\beta(f) = [x(f)/(1 - \beta)] + y(f) + \epsilon(\beta, f),$$

where $x(f)$ *is the unique solution of*

$$(I - Q(f))x = 0, \qquad Q^*(f)x = Q^*(f)r(f),$$

$y(f)$ *is the unique solution of*

$$(I - Q(f))y = r(f) - x(f), \qquad Q^*(f)y = 0,$$

and $\epsilon(\beta, f) \to 0$ *as* $\beta \to 1$.

(b) *For each* s, *denote by* $G(s, f)$ *the set of* a *for which either*

$$p(s, a)x(f) > x_s(f)$$

or

$$p(s, a)x(f) = x_s(f)$$

and

$$i(s, a) + p(s, a)y(f) > x_s(f) + y_s(f),$$

where $x_s(f)$, $y_s(f)$ *denote the* sth *coordinates of* $x(f)$, $y(f)$. *For any* g *such that* $g(s) \varepsilon G(s, f)$ *for some* s *and* $g(s) = f(s)$ *whenever* $g(s) \not\varepsilon G(s, f)$, $g > f$ *for all* β *sufficiently near* 1.

(c) *For each* s, *denote by* $E(s, f)$ *the set of* a *for which*

$$p(s, a)x(f) = x_s(f)$$

and

$$i(s, a) + p(s, a)y(f) = x_s(f) + y_s(f)$$

(always $f(s) \varepsilon E(s, f)$*). If, for each* s, $G(s, f)$ *is empty and* $E(s, f)$ *contains only the point* $f(s)$, *then* f *is optimal.*

(d) *If for each* s, $G(s, f)$ *is empty and* $g(s) \varepsilon E(s, f)$ *for all* s *implies*

$$Q^*(g)Q^*(f) = Q^*(g),$$

then f *is nearly optimal.*

(e) *For any* f_0 *for which* $G(s, f_0)$ *is empty for all* s, $x(f_0) \geqq x(g)$ *for all* g. *Denote by* F^* *the set of all* g *such that* $x(g) = x(f_0)$. *There is an* $f^* \varepsilon F^*$ *with* $y(f^*) \geqq y(g)$ *for all* $g \varepsilon F^*$. *The nearly optimal* g's *are exactly those for which* $x(g) = x(f^*)$ *and* $y(g) = y(f^*)$.

PROOF. For (a), we have

$$\begin{aligned} V_\beta(f^{(\infty)}) &= [I - \beta Q(f)]^{-1}r(f) = \sum_0^\infty \beta^n Q^n(f)r(f) \\ &= \left(\sum_0^\infty \beta^n Q^*(f) + \sum_0^\infty \beta^n(Q^n(f) - Q^*(f))\right) r(f) \\ &= \frac{Q^*(f)r(f)}{1-\beta} + H(f)r(f) + (H(\beta, f) - H(f))r(f). \end{aligned}$$

Thus (a) is established, with $x(f) = Q^*(f)r(f)$, $y(f) = H(f)r(f)$, and $\epsilon(\beta, f) = (H(\beta, f) - H(f))r(f)$. For the rest of the theorem, we simply calculate $V_\beta(g, f^{(\infty)})$, using the representation (a), and ask when, for β near 1, does this exceed $V_\beta(f^{(\infty)})$. We have

$$\begin{aligned} V_\beta(g, f^{(\infty)}) &= r(g) + \beta Q(g)V_\beta(f^{(\infty)}) \\ &= \frac{Q(g)x(f)}{1-\beta} + r(g) - Q(g)x(f) + Q(g)y(f) + \epsilon_1(\beta, f, g), \end{aligned} \tag{2}$$

where $\epsilon_1(\beta, f, g) = -(1 - \beta)Q(g)y(f) + \beta Q(g)\epsilon(\beta, f) \to 0$ as $\beta \to 1$.

We see that $g(s) \varepsilon G(s, f)$ implies that, for β near 1, the sth coordinate of $V_\beta(g, f^{(\infty)})$ exceeds that of $V_\beta(f^{(\infty)})$. Since $g(s) = f(s)$ implies equality of the sth coordinates of $V_\beta(g, f^{(\infty)})$ and $V_\beta(f^{(\infty)})$ for all β, we obtain (b) at once from Theorem 3. Similarly, the hypotheses of (c) imply that, for all β near 1,

$$V_\beta(g, f^{(\infty)}) \leqq V_\beta(f^{(\infty)})$$

(with strict inequality unless $g = f$), so that from Theorem 3 f is optimal.

For (d) we shall need

LEMMA 2. *For any $f, g \varepsilon F$ for which $g(s) \varepsilon E(s,f)$ for all s, we have $x(g) = x(f)$. If in addition $Q^*(g)Q^*(f) = Q^*(g)$, then $y(g) = y(f)$.*

PROOF OF LEMMA 2. That $g(s) \varepsilon E(s, f)$ for all s is equivalent to, writing x, y for $x(f)$, $y(f)$,

$$Q(g)x = x \tag{3}$$

and

$$r(g) + Q(g)y = x + y. \tag{4}$$

Multiplying (4) by $Q^*(g)$ yields

$$Q^*(g)r(g) = Q^*(g)x. \tag{5}$$

But (3) and (5) have the unique solution $x = x(g)$, so that $x(g) = x(f)$. Also from $Q^*(f)y = 0$ we obtain $Q^*(g)Q^*(f)y = 0$, so that, if $Q^*(g)Q^*(f) = Q^*(g)$, we obtain

$$Q^*(g)y = 0. \tag{6}$$

But, since $x = x(g)$, the unique solution of (4) and (6) is $y = y(g)$, so that $y(g) = y(f)$.

We return to (d). Let f satisfy the hypotheses of (d), and choose β so near 1 that, for any pair f_1, f_2, we have $V_\beta(f_1, f_2^{(\infty)}) \geqq V_\beta(f_2^{(\infty)})$ implies $f_1(s) \varepsilon G(s, f_1) \cup E(s, f_1)$ for all s. If our f is not β-optimal, let $f_0 = f_1, f_2, \cdots, f_k$ be a sequence of β-improvements, obtained as in Theorem 3, terminating in a β-optimal f_k. Then

$$f_{i+1}(s) \varepsilon G(s, f_i) \cup E(s, f_i)$$

for all i. We show by induction on i that $x(f_i) = x(f_0)$ and $y(f_i) = y(f_0)$. This is true for $i = 0$. If true for a given i, then, since $G(s, f)$, $E(s, f)$ depend only on $x(f)$, $y(f)$, we have $G(s, f_i)$ is empty and $E(s, f_i) = E(s, f)$. Then f, f_{i+1} satisfy the hypotheses of f, g in Lemma 2, so that $x(f_{i+1}) = x(f)$, $y(f_{i+1}) = y(f)$. Thus, writing $f(\beta)$ for the β-optimal f_k, we have

$$U(\beta) = [x(f)/(1 - \beta)] + y(f) + \epsilon(\beta, f_\beta).$$

Since

$$V_\beta(f^{(\infty)}) = [x(f)/(1 - \beta)] + y(f) + \epsilon(\beta, f),$$

we have $U(\beta) - V_\beta(f^{(\infty)}) \to 0$ as $\beta \to 1$, and $f^{(\infty)}$ is nearly optimal.

To establish (e), we obtain from (2), if $G(s, f_0)$ is empty for all s, the inequality

$$V_\beta(g, f_0^{(\infty)}) \leqq V_\beta(f_0^{(\infty)}) + \tau(\beta)\delta \qquad \text{for } \beta \text{ near } 1, \tag{7}$$

where $\tau(\beta)$ is a scalar function of β, the maximum coordinate of $\epsilon_1(\beta, f_0, g) -$

$\epsilon(\beta, f_0)$, and δ is the $S \times 1$ column vector with all coordinates unity. We have $\tau(\beta) \to 0$ as $\beta \to 1$. Denoting $L_\beta(g)$ by L, we rewrite (7) as $LV_\beta(f_0) \leqq V_\beta(f_0) + \tau(\beta)\delta$ for β near 1. We show by induction on n that, for all n

$$L^n V_\beta(f_0) \leqq V_\beta(f_0) + (1 + \beta + \cdots + \beta^{n-1})\tau(\beta)\delta \quad \text{for } \beta \text{ near 1.} \tag{8}$$

If (8) holds for a given n, we obtain, applying L,

$$\begin{aligned} L^{n+1}V_\beta(f_0) &\leqq L[\text{r.h.s. of (8)}] \\ &= r(g) + \beta Q(g) V_\beta(f_0) + \beta(1 + \beta + \cdots + \beta^{n-1})\tau(\beta)\delta, \\ &= V_\beta(g, f_0^{(\infty)}) + \beta(1 + \beta + \cdots + \beta^{n-1})\tau(\beta)\delta \\ &\leqq V_\beta(f_0) + [1 + \beta + \cdots + \beta^n]\tau(\beta)\delta, \end{aligned}$$

where the last inequality is obtained by using (7).

Thus, $L^n V_\beta(f_0) \leqq V_\beta(f_0) + [\tau(\beta)/(1 - \beta)]\delta$ for all n, so that, for all $g \varepsilon F$

$$V_\beta(g) = \lim_{n\to\infty} L^n V_\beta(f_0) \leqq V_\beta(f_0) + [\tau(\beta)/(1 - \beta)]\delta \quad \text{for } \beta \text{ near 1.} \tag{9}$$

But

$$V_\beta(g) - V_\beta(f_0) = \frac{x(g) - x(f_0)}{1 - \beta} + y(g) - y(f_0) + \epsilon(\beta, g) - \epsilon(\beta, f_0). \tag{10}$$

(9) and (10) imply $x(g) \leqq x(f_0)$.

Take any f^* which is β-optimal for a set of β's having 1 as a limit point. From (10), with $g = f^*$ we obtain $x(f^*) \geqq x(f_0)$, so that $x(f^*) = x(f_0)$. For any $g \varepsilon F^*$, we have $V_\beta(f^*) - V_\beta(g) = y(f^*) - y(g) + \epsilon(\beta, f^*) - \epsilon(\beta, g)$, so that, letting $\beta \to 1$ through a sequence for which f^* is β-optimal, we obtain $y(f^*) \geqq y(g)$ for all $g \varepsilon F^*$. The last assertion of (e) is now immediate.

Theorem 4 does not describe an algorithm which is guaranteed to lead to optimal or even near optimal policies, and which is comparable in simplicity to the algorithm described by Theorem 3 for $\beta < 1$. The algorithm is simple until we reach an f for which $G(s, f)$ is empty. At this point, if $E(s, f)$ contains for each s only the single element $f(s)$, f is optimal. If not, we know only that $x(g) \leqq x(f)$ for all g, so that we have a policy which maximizes our average return. In one case the verification of (d) is immediate. This is the case in which there is a single terminal state s^* which is certain to be reached eventually, no matter where we start or which policy we use, and which can never be left once reached. In this case for every g, $Q^*(g)$ is the matrix with every row the s^* unit vector, so that f will satisfy the hypothesis of (d) and be nearly optimal. In general, the checking of (d) is tedious and, if it fails, we are reduced to determining the set F^*, calculating $y(g)$ for each $g \varepsilon F^*$, and selecting a g for which $y(g)$ is maximal.

THEOREM 5. *There is an optimal policy which is stationary.*

PROOF. For each s and f, the sth coordinate of $V_\beta(f)$ is a rational function of β, as the representation $V = (I - \beta Q)^{-1}r$ shows. Let f^* be β-optimal for a set of β's having 1 as a limit point. Then, for every g, $V_\beta(f^*) \geqq V_\beta(g)$ for a set of β's having 1 as a limit point. Since all coordinates of $V_\beta(f^*)$ and $V_\beta(g)$ are rational functions of β,

$$V_\beta(f^*) \geqq V_\beta(g) \qquad \text{for all } \beta \text{ near } 1.$$

Since this holds for every $g \varepsilon F$, f^* is optimal.

We close with two examples.

EXAMPLE 1. *An f which satisfies the hypotheses of* (d) *of Theorem* 4, *but is not optimal.* There are two states, 1 and 2, and two actions, 1 and 2. In state 1 action 1 yields \$1, and the system remains in state 1 with probability .5 and moves to state 2 with probability .5 while action 2 yields \$2 and the system moves to state 2 with certainty. In state 2, either action yields 0 and the system remains in state 2. There are clearly only two effectively different elements of F: f:$f(1) = 1$ and g:$g(1) = 2$. We have, starting in state 1,

$$V_\beta(f^\infty) = 1 + \tfrac{1}{2}\beta + \tfrac{1}{4}\beta^2 + \cdots = 2/(2 - \beta),$$
$$V_\beta(g^\infty) = 2.$$

Thus, $U(\beta) = 2$ and $f^{(\infty)}$ is nearly optimal but not optimal. The verification that f satisfies the hypotheses of (d) of Theorem 2 is straightforward.

EXAMPLE 2. *An f for which $G(s, f)$ is empty for all s, but which is not nearly optimal.* Again there are two states, 1 and 2, and two actions, 1 and 2. In state 1, action 1 yields \$3 and the system remains in state 1 with probability .5. Action 2 yields \$6, and the system moves to state 2. In state 2, either action loses \$3 and the system remains in state 2 with probability .5 and moves to state 1 with probability .5. Again, there are only two effectively different elements of F: f:$f(1) = 1$ and g:$g(1) = 2$. Straightforward calculations yield

$$x(f) = x(g) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \qquad y(f) = \begin{pmatrix} 3 \\ -3 \end{pmatrix}, \qquad y(g) = \begin{pmatrix} 4 \\ -2 \end{pmatrix},$$

so that

$$V_\beta(g) - V_\beta(f) \to \begin{pmatrix} 1 \\ 1 \end{pmatrix} \qquad \text{as } \beta \to 1$$

and f is not nearly optimal. The verification that $G(s, f)$ is empty for each s is straightforward.

REFERENCES

[1] BELLMAN, RICHARD (1957). *Dynamic Programming.* Princeton Univ. Press.

[2] DVORETZKY, A., KIEFER, J. and WOLFOWITZ, J. (1957). The inventory problem, I and II. *Econometrica* **20** 187–222 and 450–466.

[3] HOWARD, RONALD A. (1960). *Dynamic Programming and Markov Processes.* Technology Press and Wiley, New York.

[4] KARLIN, S. (1955). The structure of dynamic programming models. *Naval Research Logistics Quart.* **2** 285–294.

[5] KEMENY, J. G. and SNELL, J. L. *Finite Markov Chains.* Van Nostrand, New York.

DISTRIBUTION LIST
CONTRACT NO. Nonr-222(53)

ORGANIZATION	NO. COPIES
Assistant Secretary of Defense for Research and Engineering Information Office Library Branch Pentagon Building Washington 25, D. C.	2
Armed Services Technical Information Agency Arlington Hall Station Arlington 12, Virginia	10
Chief of Naval Research Department of the Navy Washington 25, D. C. ATTN: Code 437, Information Systems Branch	2
Chief of Naval Operations OP-07T-12 Navy Department Washington 25, D. C.	1
Director, Naval Research Laboratory Technical Information Officer/Code 2000 Washington 25, D. C.	6
Commanding Officer, Office of Naval Research Navy #100, Fleet Post Office New York, New York	10
Commanding Officer, ONR Branch Office 346 Broadway New York 13, New York	1
Commanding Officer, ONR Branch Office 495 Summer Street Boston 10, Massachusetts	1
Bureau of Ships Department of the Navy Washington 25, D. C. ATTN: Code 607A NTDS	1
Bureau of Naval Weapons Department of the Navy Washington 25, D. C. ATTN: RAAV Avionics Division	1
Bureau of Ships Department of the Navy Washington 25, D. C. ATTN: Communications Branch Code 686	1
Naval Ordnance Laboratory White Oaks Silver Spring 19, Maryland ATTN: Technical Library	1
David Taylor Model Basin Washington 7, D. C. ATTN: Technical Library	1
Naval Electronics Laboratory San Diego 52, California ATTN: Technical Library	1
University of Illinois Control Systems Laboratory Urbana, Illinois ATTN: D. Alpert	1
Air Force Cambridge Research Laboratories Laurence G. Hanscom Field Bedford, Massachusetts ATTN: Research Library, CRX2-R	1
Technical Information Officer U. S. Army Signal Research & Development Lab Fort Monmouth, New Jersey ATTN: Data Equipment Branch	1
National Security Agency Fort George G. Meade, Maryland ATTN: R-4, Howard Campaigne	1
U. S. Naval Weapons Laboratory Dahlgren, Virginia ATTN: Head, Compution Division, G. H. Gleissner	1
National Bureau of Standards Data Processing Systems Division Room 239, Building 10 ATTN: A. K. Smilow Washington 25, D. C.	1
Aberdeen Proving Ground, BRL Aberdeen Proving Ground, Maryland ATTN: J. H. Giese, Chief Compution Lab	1
Office of Naval Research Resident Representative University of California Room 201, Building T-9 Berkeley 4, California	1
Commanding Officer ONR Branch Office John Crerar Library Building 86 East Randolph Street Chicago 1, Illinois	1
Cornell University Cognitive Systems Research Program Hollister Hall Ithaca, New York ATTN: Dr. Frank Rosenblatt	1
Communications Sciences Lab University of Michigan 180 Frieze Building Ann Arbor, Michigan ATTN: Gordon E. Peterson	1
Census Bureau Washington 25, D. C. ATTN: Office of Asst. Director for Statistical Services, Mr. J. L. McPherson	1
Stanford University Stanford, California ATTN. Electronic Laboratory, Prof. Gene Franklin	1
University of California Institute of Engineering Research Berkeley 4, California ATTN: Prof. A. J. Thomasian	1
National Science Foundation Program Director for Documentation Research Washington 25, D. C. ATTN: Helen L. Brownson	1
Wayne State University Detroit, Michigan ATTN: Dept. of Slavic Languages, Prof. Harry H. Josselson	1
University of California at Los Angeles Los Angeles 24, California ATTN: Dept. of Engineering, Prof. Gerald Estrin	1
Columbia University New York 27, New York ATTN: Dept. of Physics, Prof. L. Brillouin	1
Hebrew University Jerusalem, Israel ATTN: Prof. Y. Bar-Hillel	1
Naval Research Laboratory Washington 25, D. C. ATTN: Security Systems Code 5266, Mr. G. Abraham	1
National Physical Laboratory Teddington, Middlesex England ATTN: Dr. A. M. Uttley, Superintendent, Autonomics Division	1
Dr. Jacob Beck Harvard University Memorial Hall Cambridge 38, Massachusetts	1
George Washington University Human Resources Research Office P. O. Box 3596 Washington 7, D. C. ATTN: Dr. John Finan	1
Diamond Ordnance Fuze Laboratory Connecticut Ave. & Van Ness Street Washington 25, D. C. ORDTL-012, E. W. Channel	1
Harvard University Cambridge, Massachusetts ATTN: School of Applied Science, Dean Harvey Brook	1
Commanding Officer and Director U. S. Naval Training Device Center Port Washington Long Island, New York ATTN: Technical Library	1
Office of Naval Research Washington 25, D. C. ATTN: Code 450, Dr. R. Trumbull	1
The University of Chicago Institute for Computer Research Chicago 37, Illinois ATTN: Mr. Nicholas C. Metropolis, Director	1
U. S. Army Biological Warfare Laboratories Building 29, Room 516 Bethesda 14, Maryland ATTN: Clifford J. Maloney, Division of Biologics Standards	1
Stanford Research Institute Computer Laboratory Menlo Park, California ATTN: H. D. Crane	1
Air Force Office of Scientific Research Directorate of Information Sciences Washington 25, D. C. ATTN: Dr. Harold Wooster	1
National Bureau of Standards Washington 25, D. C. ATTN: Miss Ida Rhodes, 220 Stucco Bldg.	1
New York University New York, New York ATTN: Dr. J. H. Mulligan, Jr. Chairman, Electrical Engineering Dept.	1
Texas Technological College Lubbock, Texas ATTN: Paul G. Griffith, Dept. of Electrical Engineering	1
Prof. Frank J. Mullins c/o Bellcomm, Inc. 1737 L Street, N. W. Washington 6, D. C.	1
L. G. Hanscom Field/AF-CRL-CRRB/ Bedford, Massachusetts ATTN: Dr. H. H. Zschirnt	1
Rome Air Development Center Griffiss Air Force Base Rome, New York ATTN: Mr. Alan Barnun	1
Department of the Army Office of the Chief of Research & Development Pentagon, Room 3D442 Washington 25, D. C. ATTN: Mr. L. H. Geiger	1
Dr. George Malcolm Dyson Chemical Abstracts Ohio State University Columbus 10, Ohio	1
Royal Aircraft Establishment, Mathematics Dept. Farnborough, Hampshire, England ATTN: Mr. R. A. Fairthorne, Minister of Aviation	1
University of Pennsylvania Moore School of Electrical Engineering 200 South 33rd Street Philadelphia 4, Pennsylvania ATTN: Miss Anna Louise Campion	1
Department of the Army Office of the Asst. COFD for Intelligence Room 2B529, Pentagon Washington, D. C. ATTN: John F. Kullgren	1
Division of Automatial Data Processing/AOP/ Department of State Washington 25, D. C. ATTN: F. P. Diblasi, 19A16	1
University of Pennsylvania Mechanical Languages Projects Moore School of Electrical Engineering Philadelphia 4, Pennsylvania ATTN: Dr. Saul Gorn, Director	1
Mr. Bernard M. Fry, Deputy Head Office of Science Information Service National Science Foundation 1951 Constitution Avenue, N. W. Washington 25, D. C.	1
Harry Kesten Cornell University Dept. of Mathematics Ithaca, New York	1
Applied Physics Laboratory Johns Hopkins University 8621 Georgia Avenue Silver Spring, Maryland ATTN: Document Library	1
Bureau of Supplies and Accounts, Chief Navy Department Washington, D. C ATTN: Code W3	1
National Aeronautics & Space Administration Goddard Space Flight Center Greenbelt, Maryland ATTN: Chief, Data Systems Division, C. V. L. Smith	1
Federal Aviation Agency Bureau of Research and Development Washington 25, D. C. ATTN: RD-375, Mr. Harry Hayman	1
American Systems Incorporated 3412 Century Boulevard Inglewood, California ATTN: M. D. Adcock	1

DISTRIBUTION LIST
CONTRACT NO. Nonr-222(53)

ORGANIZATION	NO. COPIES
Commanding Officer ONR Branch Office 1030 E. Green Street Pasadena, California	1
Commanding Officer ONR Branch Office 1000 Geary Street San Francisco 9, California	1
National Bureau of Standards Washington 25, D.C. ATTN: Mr. R.D. Elbourn	1
Syracuse University Electrical Engineering Department Syracuse 10, New York ATTN: Dr. Stanford Goldman	1
Dr. W. Papian Lincoln Laboratories, MIT Lexington, Massachusetts	1
Institute for Defense Analysis Communications Research Division Von Neumann Hall Princeton, New Jersey	1
Air Force Office of Scientific Research Information Research Division Washington 25, D.C. ATTN: R.W. Swanson	1
W. A. Kozumplik, Manager Lockheed Aircraft Corporation Missiles & Space Division 3251 Hanover Street Palo Alto, California	1
Joel Levy National Bureau of Standards Far West Building, 1B Washington, D.C.	1

ORGANIZATION	NO. COPIES
The Rand Corporation 1700 Main Street Santa Monica, California ATTN: Numerical Analysis Dept., Willis H. Ware	1
Massachusetts Institute of Technology Cambridge 39, Massachusetts ATTN: Prof. John McCarthy, 26-007B	1
Sylvania Electric Systems 1100 Wehrle Drive Buffalo 21, New York ATTN: R. L. San Soucie	1
Carnegie Institute of Technology Pittsburgh, Pennsylvania ATTN: Director, Computation Center, Alan J. Perlis	1
Chief, Bureau of Naval Weapons Navy Department Washington 25, D.C ATTN: RREN	1
Electronic Systems Development Corp. 1484 E. Main Street Ventura, California ATTN: Barbara J. Lange	1
Electronics Research Laboratory University of California Berkeley 4, California ATTN: Director	1
R. Turyn Applied Research Laboratory Sylvania Electric Products, Inc. 40 Sylvan Road Waltham 54, Massachusetts	1
George Washington University Washington, D.C. ATTN: Prof. N. Grissmore	1

ORGANIZATION	NO. COPIES
Cornell Aeronautical Laboratory, Inc. P.O. Box 235 Buffalo 21, New York ATTN: Systems Requirements Dept., A.E. Murray	1
Chief, Bureau of Ships Code 671A2 Washington, D.C. ATTN: LCDR. E.B. Mahinske, USN	1
Lincoln Laboratory Massachusetts Institute of Technology Lexington 73, Massachusetts ATTN: Library	1
Maj. Gen. Casemiro Montenegro Filho, Director Centro Technico da Aeronautica/CTA Sao Jose Dos Campos Sao Paulo, Brazil	1
Professor C. L. Pekeris, Head Department of Applied Mathematics Weizmann Institute of Science Rehovoth, Israel	1
Mr. Julian H. Bigelow Institute for Advanced Study Princeton, New Jersey	1
The Mitre Corporation P. O. Box 208 Bedford, Massachusetts ATTN: Library	1
E. Tomash Ampex Computer Products P. O. Box 329 Culver City, California	1